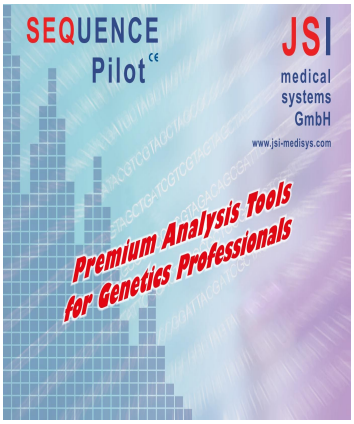
What's new

SEQUENCE Pilot CE

Version 4.1.1
(03/19/2013)



developed by

JSI medical systems GmbH
Friedhofstr. 5
77971 Kippenheim
GERMANY
phone: +49-7825/863620-0
fax: +49-7825/863620-20
email: *mail@jsi-medisys.com*

JSI medical systems Corp.
1901 Newport Blvd., Suite 350
Costa Mesa, CA 92627
USA
phone: +1-949-999-2092
fax: +1-949-999-2093
email: *mail-us@jsi-medisys.com*

web: *www.jsi-medisys.com*

(for research use only)

Table of Contents

1 All Modules....3
1.1 All tables....3
2 Modules SeqPatient, SeqC and SeqNext....3
2.1 Menu SeqNext, SeqPatient and SeqC....3
2.2 Gene Admin....4
2.3 Operation Sequence/Chip....4
2.3.1 Variation Mutation table....4
2.3.2 Report....5
3 Modules SeqPatient and SeqHLA....5
3.1 Operations Joining and Worklist....5
4 Module SeqPatient....5
4.1 Operation Joining....5
4.2 Operation Sequence....5
5 Module SeqHLA....5
5.1 Operation Sequence....5
6 Module SeqNext....6
6.1 Multiplicom MASTR assays....6
6.2 Operation ROI [master file]....6
6.2.1 All Tabs....6
6.2.2 Tab add PCR....7
6.2.2.1 Tab Gene....7
6.2.2.2 Tab Fasta....7
6.2.2.3 Tab Chromosome Range....7
6.2.3 New tab add Enrichment/Kit....8
6.3 Operation ROI Groups [master file]....9
6.4 Operation Run....9
6.4.1 Settings....9
6.4.1.1 Tab Settings....10
6.4.1.2 Tab Hide Mutation....13
6.4.1.3 Tab Quality Score....13
6.4.1.4 Tab BAM/SAM....14
6.4.1.5 Tab Trimming....14
6.4.1.6 Tab Expert Settings....15
6.5 Operations Joining/Worklist/Poollist....15
6.6 Operation Sequence....16
6.6.1 Section Order/Settings....16
6.6.2 Section Files/Info window (Files/context menu item show>Info)....16
6.6.3 Section ROI/Location table....16
6.6.4 Variation/Mutation table....16
6.6.5 Summary....17
6.6.6 Show....17
6.6.7 Electropherogram....17
6.6.8 Reads view....18
7 Module SeqNext-HLA....19
8 Module Talkmaster....19

1 All Modules

1.1 All tables

Each user can decide which table columns are shown. Therefore right click a table header. Then select *Manage table columns...*:

A dialogue opens which contains the following information:

- A list of invisible columns (on the left side)
- A list of visible columns and their width (on the right side)
- A header preview (in the lower section)

You can remove or add columns by moving them into the corresponding list. Therefore press *[Add]* or *[Remove]* or double-click an entry.
Moreover the column width can be changed in the list of visible columns by entering a number or by changing the preview below. In both cases the preview is adapted.
By pressing *[Default]* the default settings (including all columns) are applied.
By pressing *[OK]* the new settings are applied. The settings are saved for each user.

Note: In case the table is exported only visible columns are exported.

2 Modules SeqPatient, SeqC and SeqNext

2.1 Menu SeqNext, SeqPatient and SeqC

The new menu item *Import SNP DB file...* is available. Using this, SNP databases can be imported into **SEQUENCE Pilot**.

For the import of SNP database files it is required, that the genome the SNP database refers to (e.g. *hg19*) is present in the **SEQUENCE Pilot** installation (directory `Genefiles/Genome/hg19`). For instructions how to install *hg19,* please check the **User Manual SeqNext** or contact our support team.
After import: For all ROIs created using the corresponding genome (e.g. *hg19*) the SNPs are shown automatically. Genes downloaded using *Gene Admin* have to be mapped to *hg19* first, before SNPs are shown (*Gene Admin, [Extras->] >map to genome hg19*; see next chapter).

Note: In case another genome (not *hg19*) should be used for mapping, this has to be set in the `lis.ini`-file, located in the `bin`-directory of your installation: Enter the directory name of the genome behind the entry `DefaultGenomeDir=` in the section `[SeqPilot]` (e.g.: `DefaultGenomeDir=HomoSapiens`). This genome directory has to be present in the folder `Genefiles/Genome` of your installation.

Note: JSI offers an `exe`-file to import the SNP-databases "snp135" and "snp137" from **NCBI** into **SEQUENCE Pilot**. To receive this, please contact our support team.

Alternatively you can import any other SNP database doing the following:

- SNP database files can be downloaded for example using the following links.
  - **Ensembl**: in `GVF`-format - *http://www.ensembl.org/info/data/ftp/index.html*
  - **NCBI**: in `txt`-format - *http://hgdownload.cse.ucsc.edu/goldenPath/hg19/database/* e.g. `snp137.txt.gz`
- The downloaded file has to be extracted in the new directory `GeneFiles/GenomeVar` directory of your **SEQUENCE Pilot**-installation.

- To import the SNP database, do the following after selecting the menu item *Import SNP DB file...* :
  - select the genome the SNP database refers to in the field *Organism* (e.g. hg19).
  - Then press *[...]* and select the downloaded SNP database file (it locates in the `GeneFiles/GenomeVar` directory).
  - In the field *Version,* information for identification of the SNP database file such as genome, date and name of the SNP database file can be entered.
  - By pressing *[Import]* the SNPs are imported into **SEQUENCE Pilot**. This may take up to four hours. A status bar is shown during this time.

2.2 Gene Admin

The following changes are present:

- In the *isoform table* there is the new column # available to index each isoform. For **SeqPatient** only: The number listed here is the same number that is shown in the operation *Sequence* (written in parenthesis in *PCR Product/Location table*) in case several active isoforms are used to analyse a file.
- For the first active isoform it is checked, if there are several exons with the same name present. If this is the case, the user will get a warning.
- There is the new button *[renumber exon names]* available in the dialogue *Change Gene*. Using this, exons are renumbered starting from E1 (for the first exon) to EX.
- It is possible to map genes downloaded with *Gene Admin* to the genome *hg 19*. Advantages are, that the genome position is then shown in the *Variation/Mutation table*/column *Pos.* Moreover imported SNP databases (see previous chapter) are then applied. Therefore the new context menu items *[Extras ->] > map to genome hg19* > can be used*:*
  - *selected gene:* the gene selected in the *Genes table* is mapped to the genome *hg19*.
  - *all current genes*: all currently used genes (genefile selected in the *Versions* table) are mapped to the genome *hg19*. Please note, that this can take several hours, depending on how many genes are present in *Gene Admin*. You will therefore get a warning.
  - *all current + history genes*: all downloaded gene versions (including all genefiles shown in the *Version* table) are mapped to the genome *hg19*. Please note, that this can take several hours, depending on how many genes are present in *Gene Admin*. You will therefore get a warning.

  For using this item it is required that the genome *hg19* is present in the **SEQUENCE Pilot** installation (directory `Genefiles/Genome/hg19`). For instructions how to install *hg19,* please check the **User Manual SeqNext** or contact our support team.

2.3 Operation Sequence/Chip

2.3.1 Variation Mutation table

The following table columns changed:

- New column *Index*: Here the table entries are indexed.
- New column *HGVS nomenclature:* Here the *HGVS nomenclature* for each table entry is listed.
- The column *mut Ref* does not exist any more. In case the mutation is present in the internal mutation database the column *mut Entry* shows the frequency (homo/hetero) of the mutation. In case the *Nuc Name* was changed in the mutation database and differs from the entry in the column *HGVS nomenclature* it is listed in brackets.

2.3.2 Report

The entry of the column *mut Entry* (frequency of a mutation) can be listed in the report.

3 Modules SeqPatient and SeqHLA

3.1 Operations Joining and Worklist

In the field *Validation* the new search item "*required*" is available to search for orders with open requirements.

4 Module SeqPatient

4.1 Operation Joining

In case the `lis.ini` entry "forced join" is set (`JoinRFForced=yes`, section `[SeqPilot]`) the forced join to an order is done in the following cases:

- Result files with defined DNA number and SeqPrimers (sample naming) that can not be aligned to a gene are forced to join to an order (was already available in version 4.0.1).
- For result files with defined DNA number and SeqPrimers, containing less than **50** sequenced bases (or entry behind `MinRFBases=` in the `lis.ini`-file, section `[SeqPilot]`) no alignment to a gene is done. Those files are also forced to join to an order as well.

4.2 Operation Sequence

In case there is a comment present for a position/resultfile, there will be the entry * in the column *Condition* of the *Position/Resultfile table*.

5 Module SeqHLA

5.1 Operation Sequence

In the *Matching table/tab Total Result* the context menu item *show sequence specific primers* has been completely worked over. This should simplify the selection of **HLA SSP (Sequence Specific Primer) Typing Kits** from **R.O.S.E. Europe GmbH** to resolve ambiguities:

In case there are ambiguities present, the 3S-Primers, that are able to resolve the ambiguities are marked automatically. If more than one 3S-primer is able to resolve an ambiguity, all of them are provided to let the user decide which one to use. A suggested primer is marked green and required automatically.
If 3S-Primers are not sufficient, additionally exon primers are required.

The result is visualized in a window next to the table:

- it is reported, if all ambiguities can be resolved or not.
- 3S-Primers and/or additional exon primers are listed.
- resolved ambiguities are listed beneath the primers.
- unresolved ambiguities are listed.
- for result calculation 6-digit resolution is used.

6 Module SeqNext

6.1 Multiplicom MASTR assays

Multiplicom MASTR assays for copy number variation analysis:

Definition files for easy set up of ROIs for **Multiplicom MASTR assays** including copy number variation (CNV) analysis are now also available. To obtain these, please contact **Multiplicom** at *customerservice@multiplicom.com,* specifying the **MASTR assay** and **SeqNext** as target software.

After importing the ROIs, the following has to be set up for the CNV analysis (For more details see **User Manual SeqNext**):

- all ROIs belonging to the **MASTR assays** have to be grouped in the operation *ROI [master file]*
- a CNV analysis mode has to be defined for the ROI Group in the operation *Analysis Mode CNV [master file]*. Here all control ROIs have to be defined as control (*C*) and all other ROIs as targets (*T*).

6.2 Operation ROI [master file]

6.2.1 All Tabs

Note: **For old ROIs including PCR Primers, the PCR Primers have to be entered again.**

The term *amplicon part* was replaced by *amplicon*. Amplicon now refers to the reads blocks present, in case a large ROI was amplified using several primer pairs.

The field *Amplification* was renamed into *Category*. The function stays the same.

The column *Extended* was renamed into *Reference*.
The new column *Index* is available to index the table entries.

In case the location of an ROI is changed (column *Location* of the *ROI List*), defined amplicons for the ROI are adapted automatically.

ROIs can be exported or imported using the buttons *[Export]* and *[Import]* respectively. Using the export function, all ROIs listed in the *ROI List* are exported. Information such as amplicons, primers, ignored parts and pseudogenes are also exported/imported.

The fields *Amplicons* and *PCR Primers* were merged into the new field *Amplicons/PCR Primers:*

- Amplicon definition: By selecting an ROI in the list and pressing *[Add],* amplicons can be entered as done before.
- Primer definition: Primer pairs can be added for a selected ROI by pressing *[Add]*. These primer pairs then define the borders of the amplicons (primers are not included). Amplicons are therefore created automatically when a primer pair is entered.
  For primers an 5' and 3' error rate can be entered.
- *Amplicons/PCR Primers* can be inactivated by removing the mark in the column *Active*. Inactive amplicons are not used. Moreover amplicons can be defined for *both* (forward and reverse reads) or for *forward* and *reverse* reads separated, by selecting the corresponding entry in the column *Direction*.
- Moreover a `csv/txt`-file containing a list of primer pairs can be imported by pressing *[Import Primer...]*. During Import the window *Import Primer CSV/Txt Files* opens. Here you can define the file layout:

- *Separator*: Here the column separator (TAB/,/;) can be selected
- *Ignore Header*: if this is active the first line is ignored.
- *Column positions*:
  - *Sense Primer*: enter the number of the column containing the sense primer.
  - *Anti sense Primer*: enter the number of the column containing the anti sense primer.

6.2.2 Tab add PCR

6.2.2.1 Tab Gene

The direction of the gene is shown in the new column *Strand* in the *Genes* table.

In the isoform table the new columns *AA length, Start codon* and *Stop codon* are available. For each transcript the following information is listed:

- *AA length*: Number of amino acids. In case the transcript is not translated there is the entry '0' in the column *AA length*.
- *Start codon*: Exon number and position of the start codon.
- *Stop codon*: Exon number and the position of the stop codon.

In the location table the chromosomal location of each exon is shown in the column *Position*.

There is the new field *Suffix*. Optionally a suffix can be entered to be added behind the ROI name (listed in the field *Name*).

The button *[Add ->]* was renamed into *[Add 1 ->]*.

6.2.2.2 Tab Fasta

The button *[Add X ->]* is not available any longer.

The new field *Range* is available: A sequence range to be used as ROI can be entered, in case not the complete `fasta`-file should be used. The entry has the following formate: e.g. 100..200 shows, that only the bases 100 to 200 of the `fasta`-sequence are used as ROI. The number entered here is listed in the field *Location* of the *ROI list*.

6.2.2.3 Tab Chromosome Range

This tab has been worked over:

- chromosomal positions can now be entered directly into the field *Position*. In case a position is entered here, the fields *Chromosome* and *Range* are filled out automatically. Optionally the position can be entered into the fields *Chromosome* and *Range* directly.
- New button *[go]*. Press this button to start the calculation of the chromosome range.
- The button *[Add->]* in the upper part of the screen, adds the complete chromosome range as one ROI to the *ROI List*.
- Moreover single genes or exons that locate in the chromosome range can be added to the *ROI List*. Therefore you can select a gene, a transcript and exons in the corresponding tables. The tables are the same as for tab *Gene*. ROIs can be added to the *ROI List* by selecting one or several exons and pressing *[Add 1->]* in the lower part of the screen. Alternatively *[Add X->]* can be used to add several selected exons as single ROIs. Optionally a suffix to be added to the *ROI Name* can be entered in the field *Suffix*.

In the exon table all exons that are completely inside the chromosome range are marked green. Exons that do not lie within the chromosomal range are not marked. They can be added as well.

6.2.3 New tab add Enrichment/Kit

The tabs *add Enrichment* and *add Kit* were combined into the tab *add Enrichment/Kit*. Using the radio button *Enrichment* or *Kit* you can select, if you want to set up ROIs for an enrichment or a **Multiplicom MASTR assay**.

Select transcripts to build ROIs:
For enrichments as well as for **Multiplicom MASTR assays** it is now possible to select transcripts for each gene. After *[build]* is pressed, the new window *Change Gene/Transcript* opens automatically. In this window all genes defined in the enrichment-/Multiplicom-file are sorted by *Chromosomes* and *Chromosome Range*. For the selected chromosome all automatically build ROIs and used transcipts are listed in the *ROIs* table.
Above the tables the chromosome range of the selected gene and all available transcripts are shown graphically. The ROI positions defined in the file are indicated as green bars. The transcript automatically used to build the ROIs is marked red. For selection of this transcript **SeqNext** chooses the one, where most of the ROIs defined in the files cover exonic regions. In case there are several transcripts present that are equal in this matter one of them is choosen randomly. Another transcript can be selected by right-clicking on it and selecting *set to transcript*. The ROIs are changed automatically using the new transcript. When all transcripts are the desired ones, you can press *[OK]* to close the window.

Add Enrichments:

For `csv`- or `bed`-files, it can be defined in which column of the file the necessary information locates. Therefore also self-made `csv`-files containing ROI information can be loaded now. Press the button *[Columns...]* in the section *Settings*.
The new window *Import Bed Text Files* opens. The settings entered here already are the default settings for "usual" `bed`-files. They only have to be changed, according to the layout of the `bed`-file you are using. The following fields can be adapted:

- *Separator*: Here the column separator (**TAB** or **,** or **;**) can be selected.
- *Ignore Header*: If this is active the first line is regarded as header and therefore ignored.
- *Column positions*:
  - *Chr.*: Enter the number of the column containing the chromosome number.
  - *Chr Start:* Enter the number of the column containing the chromosome start.
  - *Chr End*: Enter the number of the column containing the chromosome end.
  - *Gene*: Optionally enter the number of the column containing the gene name.
  - *Transcript ID*: Optionally enter the number of the column containing the transcript ID.
  - *Comment*: Optionally enter the number of the column containing a comment.
  - *Chromosome positions include primers*: Check this option in case the chromosome positions include primers.

  Amplicon/PCR Primer:
  - *Sense Primer*: Optionally enter the number of the column containing the sense primer. It is also possible to enter a default cutting length. The number of bases entered here are then removed from each amplicon.
  - *Antisense Primer*: Optionally enter the number of the column containing the antisense primer. It is also possible to enter a default cutting length. The number of bases entered here are then removed from each amplicon.

Press *[OK]* after the fields are adapted.

Moreover the creation of ROIs out of `manifest`-files has been improved:

- Primer pairs are loaded and considered in the field *Amplicons/PCR Primers.*

6.3 Operation ROI Groups [master file]

The new buttons *[Export...], [Import...]*, *[Save as...]* and *[Import Primer...]* are available:

- *[Export...]/[Import]*: ROI Groups can be exported and imported using the buttons *[Export...]* and *[Import...]* respectively. For the ROIs included in the group all information given in the operation *ROI [master file]* (*Ignored Parts, Pseudogenes, Amplicons* and *Primers*) and CNV analysis modes are also exported and imported. Moreover you are asked if you want to export an existing profile (defined in the operation *Run*) as well. Therefore you can select a profile to export with the ROI Group.
- *[Save as...]*: Saves a copy of the selected ROI Group. You can enter a new ROI Group name and lot number.
- *[Import Primer...]*: For a selected ROI Group a file containing a list of primer pairs can be added. The primers are joined to the corresponding ROI automatically. In the operation *ROI [master file]* they are shown in the section *Amplicons/PCR Primers*. During import the window *Import Primer CSV/Txt Files opens*. Here you can define the file layout:
  - *Separator*: Here the column separator (TAB/,/;) can be selected
  - *Ignore Header*: If this is active the first line is ignored.
  - *Column positions*:
    - *Sense Primer*: enter the number of the column containing the sense primer.
    - *Antisense Primer*: enter the number of the column containing the antisense primer.

6.4 Operation Run

`sam`- and `bam`-files can be analysed now. In case you want to load `sam`- or `bam`-files, select the reference genome (e.g. *hg19*) in the field *Organism*. Moreover in the settings chosen for the run there is then new tab *bam/sam* available, to select settings for analysis of `bam`- or `sam`-files.

6.4.1 Settings

The settings in the operation *Run* changed.

`bam`- and `sam`-files can be analysed now. In case you want to load `sam`- or `bam`-files, select the reference genome (e.g. *hg19*) in the field *BAM/SAM Organism*.

The settings are sorted on the tabs *Settings*, *Hide Mutations, Quality Score, Trimming, BAM/SAM* and *Expert Settings*.

Note: The settings on tab *Settings* are very important to adapt for your analysis!

A default setting can be chosen that is selected automatically when the operation *Run* is opened. Therefore check the box *as defaul*t, before saving the profile by pressing *[Save Profile]*.

6.4.1.1 Tab Settings

Here settings can be adapted, which influence mutation calling and the sorting of mutations to the different tabs of the *Variation/Mutation table.* The settings on this tab are important to adapt for your analysis.

The first setting *[1] Single/double direction analysis* is explained in the end of this chapter, because if not switched to *off*, it can change the other settings.

Reads: The settings entered here are applied to all reads.

- *[2] Analyse/ignore region*: A region is not analysed (no variant calling) in case these settings are not fullfilled. Please note, that the settings may be changed automatically in case the single/double direction analysis is activated.
  - *Min abs coverage:* Positions with an absolute coverage below this value are written in grey, mutations are not called.

    In the combo box behind you can select, if the *Min abs coverage* has to be reached in both sequencing directions together (select *combined*), or in each direction separated (select *per dir*).

    Example: *Min abs coverage* is **20, combined**. If there is a heterozygous position with an absolute coverage of **19**, the mutation is not called.
  - *Ratio read directions:* This value is the ratio between the two read directions (forward and reverse). Positions where the ratio is not fulfilled are written in grey, mutations are not called.

    Examples:
    - *Ratio read directions* is **5%/95%**: If there is a position with **100** reads, **4** reverse and **96** forward reads, the ratio is **4%/96%**. The position is greyed out, a mutation at this position would not be called.
    - If switched to *off*, mutation calling in regions covered by one sequencing direction only is possible.

- *[3] Coverage warning > Low abs coverage warning*: The value entered here is shown as a red dotted line in the electropherogram. In case there are positions with a coverage below this value you get a warning for the ROI:
  - There is the hint *low* in the column *Coverage* of the *ROI/Location table*.
  - The graph color of the *coverage graph* (below the *location overview*) changes from grey to pink.

  The *Low abs coverage warning* can be switched off, by selecting the entry *off*. If *0* is selected, the default value defined in the `lis.ini` file will be used.

- *[4] Mutations:* The settings entered here are applied only for reads showing a mutation. If the settings are not fullfilled the mutation is not called. Please note, that the settings may be changed automatically in case the single/double direction analysis is activated.
  - *Min abs. coverage:* If the absolute coverage of the mutation is below this value, it is not called.

    In the combo box behind you can select, if the *Min abs coverage* has to be reached in both sequencing directions together (select *combined*), or in each direction separated (select *per dir*).

    Example: *Min abs coverage* is **5, combined**. If there are **4** reads present, showing the mutation, the mutation is not called.

- *Ratio read directions:* This value is the ratio between the two read directions (forward and reverse) of all reads containing the mutation. If the ratio is not fulfilled, mutations/variations are not called.

  Examples:

  - Ratio read directions is **5%/95**%: If there is a mutation detected in 100 reads and there are 4 reverse and 96 forward reads showing the mutation the ratio is not fulfilled. The mutation is not called.
  - If switched to *off*, mutation calling in regions covered by one sequencing direction only is possible.

- *Min % coverage*: If the % coverage of a mutation is below this value, the mutation will not be called.

  You can choose if the *Min % coverage* should be calculated *per direction* or *combined*:

  - *combined:* The coverages in forward and reverse sense are calculated combined.
    % coverage = total coverage of the mutation / total coverage at the position.
  - *per dir:* The coverages in forward and reverse sense are calculated separated. Both values have to reach the *Min % coverage*.
    - % forward coverage: forward coverage of the mutation / total forward coverage
    - % reverse coverage: reverse coverage of the mutation / total reverse coverage

- *[5] Mutations Sorting*: The values entered here decide, if a mutation is listed on tab *distinct, other* or *homopolymer* of the *Variation/Mutation table*. Please note, that the settings may be changed automatically in case the single/double direction analysis is activated.

  - *Distinct/Other % coverage:* If the % coverage of a mutation reaches this value, the mutation is listed on tab *distinct* of the *Variation/Mutation table*. Otherwise it is listed on tab *other.*
  - *Distinct/Homopolymer % coverage*: If the % coverage of an insertion/deletion in a homopolymer region reaches at least this value, the mutation is listed on tab *distinct* of the *Variation/Mutation table*. Otherwise it is listed on tab *homopolymer.*

    You can choose, if the *Mutations Sorting coverages* should be calculated *per direction* or *combined*:

    - *combined:* The coverages in forward and reverse sense are calculated combined.
      % coverage = total coverage of the mutation / total coverage at the position.
    - *per dir:* The coverages in forward and reverse sense are calculated separated. Both values have to reach the *Mutation sorting coverages*.
      - % forward coverage: forward coverage of the mutation / total forward coverage
      - % reverse coverage: reverse coverage of the mutation / total reverse coverage

- *[6] Homopolymer*

  - *Homopolymer region size*: The minimum number of repeats of a base to be considered as a homopolymer region.

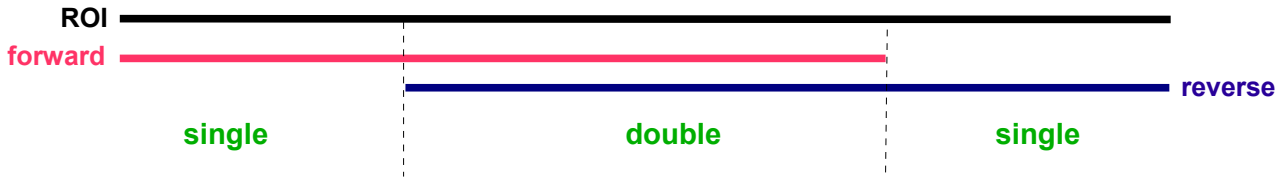
- *Reads include PCR Primer*: Please fill out the field (choose *yes* or *no*) in case you have amplicons defined, since it improves your amplicon assignment.

  - *auto*: please change this setting in case you have amplicons defined!
  - *yes*: select *yes* in case you reads include PCR primers.
  - *no:* select *no* in case your reads do not include PCR primers.

[1] Single/double direction analysis:

These settings have to be set according to the data you are analysing:

- Are all ROIs completely covered by both sequencing directions? If this is the case switch the two *Both direction coverage* settings to *off*.
- Are the ROIs partly covered by one sequencing direction only (**single**) and partly covered by both sequencing directions (**double**)?



In case you want to call mutations in the **single** regions, fill out at least on of the *Both direction min % coverage* settings:

The settings *Both directions min abs coverage* and *Both directions min % coverage* define, if the position is covered by a single direction or by double (both) directions. For **double** direction regions all other settings entered on tab *Settings* are used as they are entered. For single directions, some of the other settings are changed to the **"combined mode"**:

- *Ratio read directions* (for mutations and reads) are switched *off*
- *Min abs coverage (for reads and mutations), Min % coverage, Distinct/Other % coverage, Distinct/Homopolymer % coverage* are set to *combined.*

This way, mutations detected in single direction regions can be called. They are marked pink in the *Variation/Mutation table*, whereas mutations detected in double direction regions are not marked.

- *Both directions min abs coverage:* If the coverage is above the here entered value in both sequencing directions, we have **double** directions. If the coverage in one sequencing direction is below or equal the here entered value, we have a **single** direction.

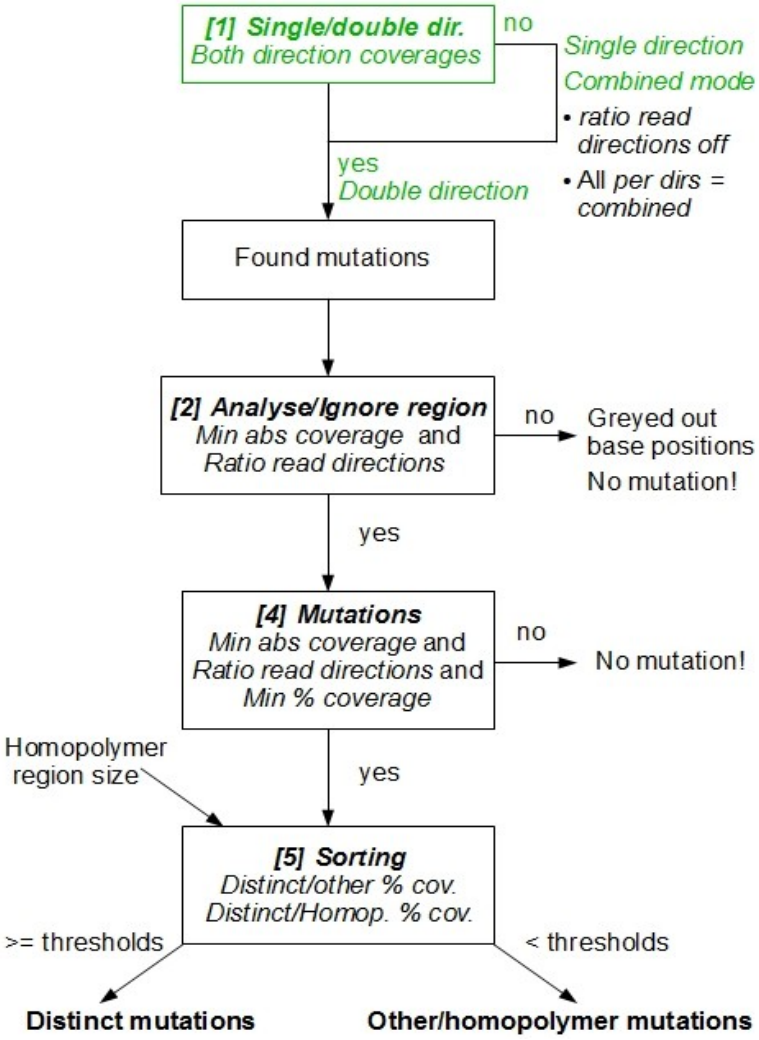
  Example: 5 is entered.

  - There are 5 forward and 100 reverse reads The region is regarded as a single direction region, the settings are changed to the combined mode.
  - There are 6 forward and 100 reverse reads. The region is regarded as a double direction region, the settings are not changed.
- *Both directions min % coverage*: If the percentage coverage is above the here entered value in both sequencing directions, we have **double** directions. If the pecentage coverage of one direction is below or equal this value, we have a **single** direction.

  Example: 5 is entered

  - There are 5 forward and 95 reverse reads (= 5 %). The region is regarded as a single direction region, the settings are changed to the combined mode.
  - There are 6 forward and 95 reverse reads (= 5,1 %). The region is regarded as a double direction region, the settings are not changed.

The flow diagram summarizes influence of the settings on the mutation calling:



6.4.1.2 Tab *Hide Mutation*

Here settings for special mutation, that should not be called can be set. If one of these items is checked, corresponding mutations are not listed in the *Variation/Mutation table:*

- *without WebRef*
- *without MutDB entry*
- *single direction reads*: mutations that are present due to the activated settings *Single/double direction analysis* (marked pink in the *Variation/Mutation table*) are not shown.
- *CNV*: mutations from CNV-analysis
- *silent mutations*
- *intronic mutations*
- *frameshift mutations*
- *splicing mutations*

6.4.1.3 Tab Quality Score

Here settings to exclude bases with bad quality from analysis can be set.

- *Quality score threshold*:

- In case there is a value entered (e.g. 15) only bases with a quality score above this value are called. Only called positions are counted to the coverage!
- The quality score filter can be switched off by selecting *Quality score threshold off*.
- In the operation *Sequence* base positions that were not called due to bad quality are shown as bases but are greyed out. In the tool tip of the forward, reverse and combined sequence, there is a new entry *Ns / Ignored,* showing how many bases at a position were ignored due to bad quality. In the coverage graph the coverage of bases that were not called due to bad quality is shown in a lighter grey.

- *Ignore reads threshold*: Here a percentage value can be entered. If the bases with bad quality in the complete read (including primers/adapters) exceed this value, the read is ignored. Bad qualtiy means, that the quality is below or equal the *Quality score threshold* setting, or that Ns were called by the sequencer.
- *Low Quality score coverage warning:* Here a threshold to get a warning for ROIs with a bad quality score can be set: In case bases with low quality exceed the here entered percentage value at a position in the ROI, the warning *low quality* is shown in the column *Coverage* of the *ROI/Location table*. You can jump to positions with low quality using the jumper *low quality* in the section *Show*.

6.4.1.4 Tab BAM/SAM

`sam`- and `bam`- files can be analysed now. Here you can choose the settings used:

- *use BAM/SAM mapping*: If this setting is active, the *BAM/SAM* mapping is used. **SEQUENCE Pilot** does not do an own mapping.
- *use BAM/SAM alignment*: If this setting is active, the *BAM/SAM* alignment is used. **SEQUENCE Pilot** does not do an own alignment.

6.4.1.5 Tab Trimming

Here settings to trim adapters or to automatically remove sequences at the ends of the reads can be entered. The settings adaptable here, were available in the previous version already:

- *Adapter*: Here adapter sequences can be entered to trim or discard reads (e.g. for **Haloplex** or **Fluidigm**). The sequence entered here is searched for (also in reverse complement in all reads): Enter an adapter sequence in 5'->3' direction. The following fields can be edited for each entered adapter:
  - Position
    - *auto*: It is decided automatically, if the adaptor locates at the 5' or 3' end.
    - *5'*: Adapter can be found at the 5' end only.
    - *3'*: Adapter can be found at the 3' end only.
  - *Error rate*: here a percentage value can be entered as error rate (wrong bases, that the adapter can contain).
  - *Overlap*: Here the minimum number of adapter bases that must overlap with the read can be entered. Example: overlap is *3*. There must be at least *3* adapter bases found in the read.
- *Remove ends*: This function allows to automatically remove sequences at the ends of the reads (e.g. adapters) that do not match to the reference: The user can decide to remove bases at the 5' and/or 3'-end by activating the box. For removing ends a number of bases beginning from the end of the read to the number given in the field *Distance* is checked.
  - In case the are no mismatches nothing is removed.
  - In case there are mismatches it is checked if sequences start to match within the *Distance*

area.

- In case a match is found, the matching sequence must have at least the length entered in the field *Blocklength*. It can include as many mismatches as entered in the box *Max. mismatches*. The bases before (for 5') or behind (for 3') the first matching base are then removed.
- In case there is no matching block found nothing is removed.

- *Remove bases*: Removes bases at the beginning and/or end of each read. Adaptors are removed prior to that.

6.4.1.6 Tab Expert Settings

Further settings concerning read processing and mutation detection can be adapted on tab *Expert Settings*. On this tab tool tips are available, explaining the settings.

- *Base calling:*
  - *Genome set:* If set to haploid, the basecalling is expected to be homozygous.
  - *Unique reads only*: If checked, the coverage of identical reads is set to 1.
- *Read processing:*
  - *Alignment evaluation:*
    - *Skip evaluation*: If checked, no alignment evaluation is done. The filters below are not used.
    - *Max mismatches*: Filter for mismatches a read can contain compared to the reference. The higher the number entered here is, the more mismatches are accepted. In case there are too many mismatches the read is discarded.
    - *Min matching bases:* Percentage of read bases that have to match to the reference. In case less bases match, the read is discarded.
    - *Keep strong consensus*: The percentage of consecutive bases that have to match to the reference without a mismatch between them. If this value is reached, the settings *Max mismatches* and *Min Matching bases* are overruled and the read is aligned. This filter is only applied if the read length is above 100.
  - *Compl. reads only*: If checked, reads that do not cover the complete amplicon are discarded.
  - *Barcode 5'+3'*: Choose this setting in case barcodes have to be present at both ends of the reads.
  - *ignore paired end info*: If checked, paired end information is not used.
- *Mutation table:*
  - *Warning:* Percentage of bases of a read that must be not aligned to the reference, before the warning *W* appears in the *Variation/Mutation table*.
  - *Indel gap SNP to SNP:* The maximum number of bases between two base changes to consider those two base changes as an InDel.
  - *Indel gap SNP to InDel:* The maximum number of bases between a base change and an InDel to consider them as an InDel.

6.5 Operations Joining/Worklist/Poollist

In the search field *Gene*, all genes with defined ROIs (in the operation *ROI [master file]*) are available. In older versions only genes, that were defined in the `lis.ini`-file were displayed.

If a gene is selected in the search field *Gene*, only orders including this gene are shown in the *Lower table*. When orders are opened the sequences of the selected gene is shown automatically. When

you jump to previous or next order using the buttons *[Previous]* or *[Next]* respectively the gene is selected as well.

6.6 Operation Sequence

6.6.1 Section Order/Settings

MIDs are shown behind the *Barcode* in parenthesis. This also works for old orders (analysed with version 4.0.1 or previous).

6.6.2 Section Files/Info window (*Files*/context menu item *show>Info*)

MIDs are shown behind the *Barcode/BarcodeExt* in parenthesis. Old orders (analysed with version 4.0.1 or previous) have to be recalculated to show MIDs.

All settings used for the Run are listed.

The section *ROI*, showing the number of aligned reads is not available any more in the *Info* window. These informations are now shown using the context menu item *show* > *info* in the *ROI/Location table* or using the *Summary* button.

6.6.3 Section ROI/Location table

The new context menu item *show* > *info* is available now. Here the number of assigned and aligned reads for the selected ROI are shown. Moreover the number of trimmed reads is displayed.

The *ROI Info window* (context menu item *ROI info*) has two tabs now: *Future analysis* and *Current analysis*.
On tab *Current analysis* information about the ROI used for the current run are shown.
ROIs can be changed on tab *Future analysis* only. In case the ROIs are changed, the ROIs that will be used for future runs are shown on tab *Future analysis*. Changes, such as new *Amplicons, Ignored Parts* or *Ignored Sequences* are highlighted. After doing a recalculate the highlighting disappears and the new ROI information are also shown on tab *Current analysis*.

In the column *Coverage* there is the new warning *low quality* available. This is shown in case the setting *Low Quality score coverage warning* (operation *Run*) is exceeded. You can jump to positions with low quality using the jumper *low quality* in the *Show* window.

6.6.4 Variation/Mutation table

For the module **SeqNext** it is possible to set **SeqPatient** requirements for entries of the *Variation/Mutation table*. Therefore the new context menu item *require SeqPatient primer* is available. This item opens a new window where *SeqPrimers* to be required are proposed automatically. In case a requirement is set, there is the entry *Q* in the column *Condition* of the *Variation/Mutation table*. Moreover a **SeqPatient** order is created automatically including the required genes/positions. The name of the **SeqPatient** order corresponds to the name of the **SeqNext** order. You can jump between **SeqNext** and **SeqPatient** by using *[Extras->]* > *Jump to Sequence*.

Moreover it can be entered if a mutation was confirmed with **SeqPatient**. Therefore the new context menu item *confirmation* is available to show, if a mutation was confirmed or rejected with Sanger sequencing. In case a mutation is confirmed, there is a + in the column Condition, in case it is rejected there is a -. The confirmation can be reset by selecting *confirmation* > *reset*.

There is the new column *Index* to number mutations starting with 1.

There is the new tab *filter* in the *Variation/Mutation table*: This tab has only entries in case the filter function of the *Variation/Mutation table* is used. It then lists all mutations that do not pass the filter settings and are therefore filtered from the other tabs of the *Variation/Mutation table.*

Mutations can be filtered using the new the context menu item *filter ...*. This serves to get an impression on how the *Variation/Mutation table* changes, when other settings are used. This function helps to optimize you settings used for the run.
In the *Filter* window the *Settings* used in the operation *Run,* on *tab Settings* and *tab Hide Mutations* can be changed. Not all settings can be changed, it is only possible to change settings that reduce the number of mutations listed in the *Variation/Mutation table*. The section *Variations* shows the number of variations/mutations, listed on the five tabs of the *Variation/Mutation table,* when the settings shown above are applied.
When *[OK]* is pressed, the new settings are applied. All variations/mutations listed in the *Variation/Mutation table* are then sorted to the tabs, regarding the new settings. Tab *Filter* now lists all variation/mutations that do not pass the filter settings and would not be detected, in case the run was recalculated using the filter settings.
When the operation Sequence is left the filter is reset and all variations/mutations are re-sorted to the tabs they have been listed on before.

In the *Filter* window a setting can be saved as a profile by pressing the button *[Save Profile]*. This profile can then be used to recalculate the file to apply the new settings.

6.6.5 Summary

The seach function is not available any more.

On all tabs the new column *Index* is available. In this columns the entries are numbered starting from 1.

The new tab *ROI* is available, which shows information about the ROI. This tab was named *Amplicon* in older versions. On tab *ROI* the *average coverage fwd/rev* is listed for the complete ROI. Moreover the number of assigned reads, trimmed reads and aligned reads is listed in the corresponding columns.

The tab *Amplicon* now corresponds to the defined amplicons in case an ROI was sequenced using several amplicons. There are only entries present for ROIs with defined amplicons. Here the a*verage coverage fwd/rev* for each amplicon is shown. Moreover the number of aligned reads is listed in the corresponding column.

6.6.6 Show

There is the new jumper *low quality* available to jump to positions with low quality. Those positions are present in case the setting *Low Quality score coverage warning* (operation *Run*) is exceeded. In this case there is also the warning *low quality* in the column *Coverage* of the *ROIs/Locations table.*

6.6.7 Electropherogram

The user can decide to show either one electropherogram for the combined sequence or two electropherograms for forward and reverse sequence separated. Therefore the new combo-box *combined/fwd+rev* is available below the *count mode* combo-box.

Min coverage line: For orders analysed with software version 3.5.2 and older, the value for the *Min coverage line* is "*0*", therefore the default settings of the `lis.ini` file is used as *Min coverage line*.

The context menu item *show fragments view...* available in combined, forward and reverse sequence was renamed into *show reads...*.

In the in the main view as well as in the *reads view* the count mode *amplicon* was renamed into *ROI abs*.

Ignored bases:

- Bases in *forward, reverse* and *combined sequence,* that are excluded from analysis are written in grey. Bases are excluded from analysis in case the settings *Ratio reads directions*

and *Min abs coverage* for the *Reads* entered in operation *Run* are not fulfilled.

- Bases in the reads sequences, that are excluded from analysis are greyed out. Bases can be excluded from analysis due to bad quality scores, depending on your settings in the operation *Run*/tab *Quality Score*. In the tool tip of the forward, reverse and combined sequence, there is a new entry *Ns/Ignored*, showing how many bases at a position were ignored due to bad quality. In the coverage graph the coverage of bases that were not called due to bad quality is shown in a lighter grey.

There is several new information available in the tool tip of forward, reverse and combined sequence:

- *Coverage*: This is the number of called bases at a position.
- *Reads*: This is the number of reads covering a position.
- *Ns/Ignored*: This is the number of Ns and ignored bases (due to bad quality) at a position.

6.6.8 Reads view

New field *quality*: if this field is activated, quality scores are shown as a color code below the bases. The item replaces the old item *bases + quality* in the combo box. The following colors are used by default:

- quality score 1-10=Dark Red
- quality score 11-20=Red
- quality score 21-30=Yellow
- quality score 31-40=Light Green
- quality score 41-99=Dark Green

Original reads:

- The original length of the reads is shown when the field *org. reads* is checked. The minimum and maximum number of bases, that were removed left and right is indicated in color.
- To see the sequence of the removed bases, the new context menu item *show original reads...* is available. When this item is selected a new *reads view* opens, showing original reads. Bases that were removed at the left and the right a highlighted.
- The read identifier is shown in the tool tip of the read, when the box *org. reads* is checked. Here the file and the line which the read can be found is shown.
- The read identifier is also shown as a tooltip when the new context menu *show original reads...* is used. For paired end sequencing data the paired end key is shown in this tooltip as well.

New button *[Search]* is available: Using this, you can either search for a sequence in the reads or for read by giving the line number of the Next Generation Sequencing file. The *Search* window opens when *[Search]* is pressed:

- To search for a sequence:
  - Enter a sequence to be searched for in the field *Sequence*
  - After pressing *[Search]* all reads containing this sequence are listed in a new window. The sequence that was searched for is highlighted.
- To search for a read listed in the Next Generation Sequencing File:
  - Select the file in the first field of the section *Line*.
  - Enter the line number in the second field of the section *Line*.
  - When *Show read group* is active not only the read present in the entered line, but all identical reads are listed.
  - After pressing *[Search]* the read(s) is (are) listecd in a new window. The line information is available in the tooltip of the read, when *org. read* is *active*.

7 Module SeqNext-HLA

The module **SeqHLA-454** was renamed into **SeqNext-HLA**.

8 Module Talkmaster

Orders with open requirements can be exported in the *Joining* and *Worklist*. To search for orders with open requirements the new search item "*required*" is available in the field *Validation* of the *Joining* and *Worklist*.

A new tag *Confirmation* was added to the **TalkMaster** export. Using this, "confirmed" or "rejected" can be exported